


METHOD AND DEVICE FOR DETERMINING PITCH WAVEFORM SEGMENTATION REFERENCE POSITION

Patent number: JP10260697
Publication date: 1998-09-29
Inventor: NISHIMURA HIROFUMI; MINOWA TOSHIMITSU; ARAI YASUHIKO
Applicant: MATSUSHITA ELECTRIC IND CO LTD
Classification:
- international: G10L9/14
- european:
Application number: JP19970063351 19970317
Priority number(s):

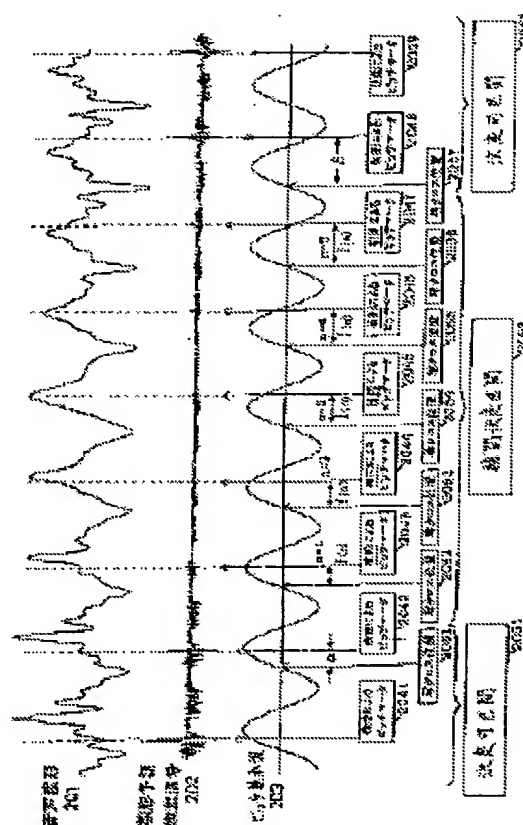
Also published as:

 JP10260697 (

Abstract of JP10260697

PROBLEM TO BE SOLVED: To determine an accurate pitch mark even in a section wherein a pitch mark can not be determined with a linear predictive residue signal.

SOLUTION: The time interval *a between the final pitch mark 20 of a determination section 2051 and the zero-cross position 2061 of a pitch basic wave 203 right before the position is found. The time interval *b between the 1st pitch mark 2048 of the determination section 2053 and the zero-cross position 2067 of the pitch basic wave 203 right before the position is found. The wave number of pitch basic waves in an interpolation determination section 2052 is find by counting zero crossing where the pitch basic wave 203 varies from negative to positive. Position which are f (n) precedent to respective zero-cross positions of all negative-to-positive varying zero-cross positions 2062 to 2066 in the interpolation determination section 2052 are regarded as pitch marks 2043 to 2047 found through interpolation. Even in a section wherein a pitch mark can not be determined from the linear residue signal, an accurate pitch mark can be determined through the interpolation.



Data supplied from the esp@cenet database - Worldwide

BEST AVAILABLE COPY

(19) 日本国特許庁 (J P)

(12) 公開特許公報 (A)

(11) 特許出願公開番号

特開平10-260697

(43) 公開日 平成10年(1998) 9月29日

(51) Int.Cl.⁶

G 1 0 L 9/14

識別記号

F I

G 1 0 L 9/14

B

審査請求 未請求 請求項の数 7 O L (全 12 頁)

(21) 出願番号 特願平9-63351

(22) 出願日 平成9年(1997) 3月17日

(71) 出願人 000005821

松下電器産業株式会社

大阪府門真市大字門真1006番地

(72) 発明者 西村 洋文

神奈川県横浜市港北区綱島東四丁目3番1号 松下通信工業株式会社内

(72) 発明者 養輪 利光

神奈川県横浜市港北区綱島東四丁目3番1号 松下通信工業株式会社内

(72) 発明者 新居 康彦

神奈川県横浜市港北区綱島東四丁目3番1号 松下通信工業株式会社内

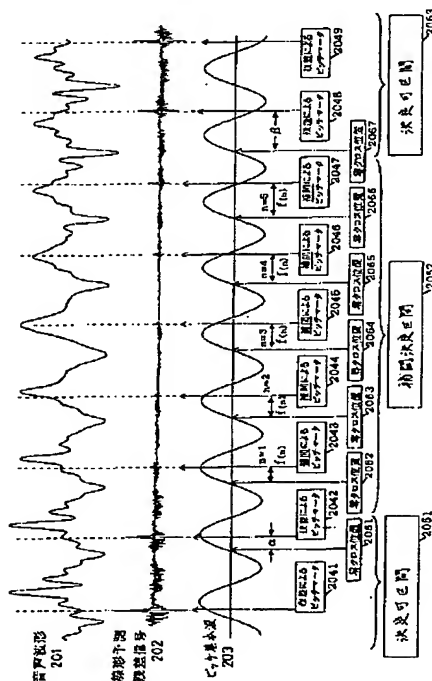
(74) 代理人 弁理士 松村 博

(54) 【発明の名称】 ピッチ波形切り出し基準位置決定方法とその装置

(57) 【要約】

【課題】 線形予測残差信号からピッチマークを決定することができない区間においても、正確なピッチマークを決定する。

【解決手段】 決定可区間2051の最後のピッチマーク2042とその位置の直前にあるピッチ基本波203の零クロス位置2061までの時間間隔 α を求める。決定可区間2053の最初のピッチマーク2048とその位置の直前にあるピッチ基本波203の零クロス位置2067までの時間間隔 β を求める。補間決定区間2052のピッチ基本波の波数をピッチ基本波203の負から正へ変化する零クロス位置を数えて求める。この補間決定区間2052の全ての負から正へ変化する零クロス位置2062~2066の各零クロス位置から $f(n)$ だけ直後の位置を補間処理により求めたピッチマーク2043~2047とする。線形予測残差信号からピッチマークを決定できない区間でも、補間処理により正確なピッチマークを決定することができる。



【特許請求の範囲】

【請求項1】 音声波形から線形予測残差信号と、ピッチフィルタを通したピッチ基本波を求め、前記線形予測残差信号からピッチ波形を切り出す正確な基準位置を決定できる区間であるかどうかを判断する工程により、正確な基準位置を決定できないと判断された区間では、前記ピッチ基本波の零クロス位置と、前記ピッチ基本波より求めたピッチの波数と、前記線形予測残差信号により正確な基準位置が決定できると判断された区間で決定した基準位置とを用いて、当該区間の基準位置を補間処理し有声区間全体で正確な基準位置を決定することを特徴としたピッチ波形切り出し基準位置決定方法。

【請求項2】 線形予測残差信号から正確な基準位置を決定できる区間であるかどうかを判断する工程において、判断する位置毎にピッチ基本波のピッチ周期程度の範囲で線形予測残差信号の局所的な最大値を求め、単語や音節や有声区間などの比較的広い範囲で求めた線形予測残差信号の広範囲の最大値に対して前記局所的な最大値がどのくらいの比率であるのかを求め、該比率を判断基準にして前記区間を判断することを特徴とする請求項1記載のピッチ波形切り出し基準位置決定方法。

【請求項3】 線形予測残差信号から正確な基準位置を決定できる区間であるかどうかを判断する工程において、判断する位置毎に該位置の音声波形の形状を調べ、前記音声波形が線形予測残差信号にピークの現れにくい正弦波のような形状の場合には、線形予測残差信号からは正確な基準位置を決定できない区間であると判断することを特徴とする請求項1記載のピッチ波形切り出し基準位置決定方法。

【請求項4】 線形予測残差信号から正確な基準位置を決定できる区間であるかどうかを判断する工程において、線形予測残差信号から正確な基準位置を決定できないと判断される区間の長さに対して最短処理の区間長を設けて、補間処理に必要な区間長に延長してから基準位置の補間処理を行うことを特徴とする請求項1記載のピッチ波形切り出し基準位置決定方法。

【請求項5】 音声波形から線形予測残差信号と、ピッチフィルタを通したピッチ基本波を求め、前記線形予測残差信号からピッチ波形を切り出す正確な基準位置を決定できる区間であるかどうかを判断する工程により、正確な基準位置を決定できると判断された区間では、短区間変形自己相関係数の最大位置により基準位置を決定し、正確な基準位置を決定できないと判断された区間では、該区間に最も近くにある前記短区間変形自己相関係数の最大位置により決定した基準位置と、該基準位置の直前にあるピッチ基本波の零クロス位置との間隔を求め、当該区間内にある全ての零クロス位置に対して、該零クロス位置から前記間隔の直後の位置に基準位置を決定するようにしたことを特徴とするピッチ波形切り出し基準位置決定方法。

【請求項6】 音声波形の有声／無声を判定する有声／無声判定部と、音声波形から線形予測残差信号を求める線形予測逆フィルタと、音声波形からピッチ基本波を求めるピッチフィルタと、前記線形予測残差信号から正確な基準位置を決定できる区間であるのか、補間処理により基準位置を決定する区間であるのかを分類する区間分類部と、該区間分類部で線形予測残差信号から正確な基準位置を決定できると判断された区間に対して、前記線形予測残差信号から基準位置を決定する基準位置決定部と、前記区間分類部で補間処理により基準位置を決定すると判断された区間に対して補間処理を行う補間部を備えたことを特徴とするピッチ波形切り出し基準位置決定装置。

【請求項7】 前記補間部は、ピッチの波数を抽出するピッチ波数抽出部と、ピッチ基本波の零クロス位置を抽出する零クロス位置抽出部と、前記ピッチの波数と前記零クロス位置と線形予測残差信号より決定した正確な基準位置とから、補間処理により基準位置を決定すると判断された区間に対して、基準位置を補間処理する基準位置線形補間部を備えたことを特徴とする請求項6記載のピッチ波形切り出し基準位置決定装置。

【発明の詳細な説明】

【0001】

【発明の属する技術分野】本発明は、音声合成する音声合成方法およびその装置において、特に窓関数によりピッチ波形を切り出し、再び重ねて音声の合成をするピッチ波形切り出し基準位置決定方法とその装置に関するものである。

【0002】

【従来の技術】従来、音声波形からピッチ波形を切り出すには、特開平8-234794号公報に記載されているように、有声区間毎に線形予測残差信号の最大値の位置を求め、これを基準位置(以下、ピッチマークという)を求める開始位置とし、隣接するピッチマークを短区間変形自己相関係数の最大位置により決定し、有声区間の全てのピッチ波形のピッチマークを順次求めるといったものである。

【0003】

【発明が解決しようとする課題】しかしながら、有声部分の一部で線形予測残差信号の振幅が極めて小さくなる場合や線形予測残差信号に顕著なピークがない場合といったようなことがあると、ピッチマークが正しく決定できないという問題があった。

【0004】本発明は、前記従来技術の問題を解決することに指向するものであり、線形予測残差信号からピッチマークを決定することができない区間においても、正確なピッチマークを得ることができるピッチ波形切り出し基準位置決定方法とその装置を提供することを目的とする。

【0005】

【課題を解決するための手段】この目的を達成するために、本発明に係るピッチ波形切り出し基準位置決定方法とその装置は、音声波形の有声/無声を判定する有声/無声判定部と、音声波形から線形予測残差信号を求める線形予測逆フィルタと、音声波形からピッチ基本波を求めるピッチフィルタと、線形予測残差信号から正確な基準位置を決定できる区間であるのか、補間処理により基準位置を決定する区間であるのかを分類する区間分類部と、区間分類部で線形予測残差信号から正確な基準位置を決定できると判断された区間に対して、線形予測残差信号から基準位置を決定する基準位置決定部と、区間分類部で補間処理により基準位置を決定すると判断された区間に対して補間処理を行う補間部を備えたことを特徴とする。

【0006】また、補間部は、ピッチの波数を抽出するピッチ波数抽出部と、ピッチ基本波の零クロス位置を抽出する零クロス位置抽出部と、ピッチの波数と零クロス位置と線形予測残差信号より決定した正確な基準位置とから、補間処理により基準位置を決定すると判断された区間に対して、基準位置を線形補間する基準位置線形補間部とを備えるように構成したものである。

【0007】前記構成によれば、正確なピッチマークを決定できないと判断された区間では、ピッチ基本波の零クロス位置と、ピッチの波数と、線形予測残差信号により決定されたピッチマークとを用いて、区間のピッチマークを補間処理し有声区間全体で正確なピッチマークを決定し、そのピッチマークを用いて、ピッチ波形を抽出し、再び重畳して音声を合成すれば、高品質の合成音声を得ることができる。

【0008】また、線形予測残差信号からピッチマークを決定できる区間であるのかどうかの判断において、判断する位置毎にピッチ周期程度の範囲で線形予測残差信号の局所的な最大値を求め、比較的広い範囲で求めた線形予測残差信号の広範囲の最大値と局所的な最大値の比率を求め、この比率により判断することで、線形予測残差信号の顕著なピークはあるがパワーが小さい場合を適確に分類することができる。

【0009】また、線形予測残差信号からピッチマークを決定できる区間であるのかどうかの判断において、判断する位置毎にその付近の音声波形の形状を調べ、音声波形が線形予測残差信号にピークの現れにくい正弦波のような形状のとき、線形予測残差信号からは正確なピッチマークを決定できない区間と判断することで、線形予測残差信号では判断が困難である区間を適切に分類することができる。

【0010】また、線形予測残差信号からピッチマークを決定できる区間であるのかどうかの判断において、線形予測残差信号からピッチマークを決定できないと判断される区間の長さに対して最短処理の区間長を設けて、補間処理に必要な区間長に延長してから基準位置の補間

処理を行うことで、補間処理によりピッチマークを決定区間が非常に短いときに起こるピッチマークの間隔の急激な変化を抑えることができる。

【0011】また、正確なピッチマークを決定できないと判断された区間では、その区間に最も近くにある決定されたピッチマークと、その直前にあるピッチ基本波の零クロス位置との間隔を求め、その区間内にある全ての零クロス位置に対して、その零クロス位置から求めた間隔の直後の位置にピッチマークを決定し、区間の一方の端で有声区間が終了する場合でも正確なピッチマークを決定することができる。

【0012】

【発明の実施の形態】以下、図面を参照して本発明における実施の形態を詳細に説明する。図1は本発明の実施の形態1におけるピッチ波形切り出しを行うピッチマーク決定方法の概念図を示したものである。図1において、101は音声波形、102は線形予測残差信号、103はピッチ基本波、1051、1053、1054は線形予測残差信号から正確なピッチマークを決定できると判断された決定可区間、1052は線形予測残差信号からでは正確なピッチマークを決定できないと判断された決定不可区間である。

【0013】まず、音声波形101から線形予測残差信号102を求め、これより、正確なピッチマークを決定できるかどうかの判定を行う。この判定方法は、例えば、判別する位置付近における線形予測残差信号の短区間平均パワーが、ある閾値以上の場合は線形予測残差信号により正確なピッチマークを決定できる区間であると判断し、そうでない部分は線形予測残差信号からは正確なピッチマークを決定できない区間と判断する。

【0014】決定可区間1051、1053、1054は従来の方法により正確なピッチマークを決定することができる。つまり、決定可区間1051の中で線形予測残差信号の最大値の位置を求め、この位置をこの部分のピッチマークとすると共に、決定可区間1051のピッチマーク決定開始位置とする。このピッチマーク決定開始位置を基準としてピッチ周期だけ負の時刻の位置で短区間変形自己相関係数の最大値の位置を求め、この位置をピッチマーク決定開始位置より一つ前に隣接するピッチマークにする。そして、このピッチマーク決定開始位置より一つ前に隣接するピッチマークの位置を新たな基準として、さらに負の時刻の方向に隣接するピッチマークを同様に決定する。この作業を繰り返し、決定可区間1051内でピッチマーク決定開始位置から負の時刻の方向にあるピッチマークを全て決定する。次に、決定可区間1051内でピッチマーク決定開始位置から正の時刻の方向にある全てのピッチマークをピッチマーク決定開始位置より負の時刻の方向のピッチマークを決定したのと同様の手順で決定する。決定可区間1053、1054についても決定可区間1051と同様にしてピッチマークを決定する。

【0015】次に決定不可区間1052のピッチマークを決

定する。決定不可区間1052のように線形予測残差信号のパワーが小さすぎたり、顕著なピークが現れない場合には線形予測残差信号からは正確なピッチマークを決定することができない。このような区間はピッチ基本波を利用してピッチマークの補間処理を行う。

【0016】図2はピッチ基本波を利用したピッチマークの補間方法の概念図を示したものである。図2において、201は音声波形、202は線形予測残差信号、203はピッチ基本波、2041、2042、2048、2049は線形予測残差信号から決定したピッチマーク、2043～2047は補間処理により決定されたピッチマーク、2051、2053は線形予測残差信号から正確なピッチマークを決定できると判断した決定可区間、2052は補間処理によりピッチマークの決定を行う補間決定区間、2061～2067はピッチ基本波203の負から正へ変化する零クロス位置である。

【0017】まず、決定可区間2051の最後のピッチマーク2042とその位置の直前にあるピッチ基本波203の零クロス位置2061までの時間間隔 α を求める。同様に、決定可区間2053の最初のピッチマーク2048とその位置の直前にあるピッチ基本波203の零クロス位置2067までの時間間隔 β を求める。次に、補間決定区間2052のピッチ基本波の波数を求める。波数を求めるにはピッチ基本波203の負から正へ変化する零クロスを数えればよく、図2の例に示す補間決定区間2052には5個の波数がある。そして、この補間決定区間2052の全ての負から正へ変化する零クロス位置2062～2066の各零クロス位置から(数1)に示す $f(n)$ だけ直後の位置を補間処理により求めたピッチマーク2043～2047とする。

【0018】

【数1】 $f(n) = \alpha + (\beta - \alpha) n / (N + 1)$

N: 波数(この例では5個)

n: 補間によりピッチマークを決定する

区間の中でのピッチマークの順番

このように、本実施の形態1によれば、線形予測残差信号から正確なピッチマークを決定できない区間でも、補間処理により正確なピッチマークを決定することができる。したがって、このように決定したピッチマークを用いて、ピッチ波形を抽出し、再び重畳して音声合成すれば、極めて高品質の合成音声を得ることができる。

【0019】図3は本発明の実施の形態2における線形予測残差信号のピッチマークを決定する区間を分類する方法の概念図を示したものである。図3において、301は線形予測残差信号、302、303は線形予測残差信号から正確なピッチマークを決定できるかどうかを判断しようとしている決定判断位置である。

【0020】まず、単語や音節や有声区間といった比較的広い範囲で線形予測残差信号の最大値を求め、これを広い範囲の最大値として、図3に示す広範囲の最大値304とする。この広範囲の最大値304に0.3～0.5程度の係数 γ を掛けて、この値を閾値305とする。

【0021】次に、全ての決定判断位置について線形予測残差信号から正確なピッチマークを決定できるかどうかの判断を行う。ここでは、図3に示す決定判断位置302、303の2箇所の位置について具体的に説明する。まず、決定判断位置302と決定判断位置303のそれぞれの位置を中心としてピッチ基本波のピッチ周期より少し広い範囲で線形予測残差信号301の局所的な最大値 $\max 1$ 、 $\max 2$ をそれぞれ求める。この局所的な最大値 $\max 1$ 、 $\max 2$ と先に求めた閾値305の大きさを比較して、図3に示す例では、局所的な最大値 $\max 1$ が閾値305より小さい場合に線形予測残差信号からは正確なピッチマークを決定できない決定判断位置302であると判断し、局所的な最大値 $\max 2$ が閾値305より大きい場合には線形予測残差信号から正確なピッチマークを決定できる決定判断位置303であると判断する。

【0022】単純に線形予測残差信号の短区間平均パワーを判断基準にすると、線形予測残差信号に顕著なピークが十分に現れているがパワーが小さい場合において、線形予測残差信号から正確なピッチマークを決定できない位置であると判断してしまうが、局所的な最大値を判断基準にすれば、適切な判断をすることができる。

【0023】このように、局所的な最大値を基準にして判断を行うことにより、適切に区間を分類することができるので、それぞれのピッチマークを適切な方法でより正確に決定することができる。

【0024】図4は本発明の実施の形態3における線形予測残差信号のピッチマークを決定する区間を分類する方法の概念図を示したものである。図4において、401は音声波形、402、403、404は線形予測残差信号から正確なピッチマークを決定できるかどうかを判断しようとしている決定判断位置、411、413、415は位置402、403、404の直前の零クロス位置、412、414、416は位置402、403、404の直後の零クロス位置である。

【0025】一般的に音声波形が正弦波のような波形の場合、顕著なピークはあまり現れない。したがって、音声波形が正弦波のような波形の場合には線形予測残差信号からは正確なピッチマークを決定することが困難である。またこのような場合には、線形予測残差信号から正確なピッチマークを決定できるかどうかを判断するのではなく、元の音声波形から判断をした方が容易、かつ適切な判断をすることができる。そこで、音声波形が正弦波のような波形である場合については、線形予測残差信号からは正確なピッチマークを決定できない区間であると判断するようにする。

【0026】次に、その判断方法を図4に示す決定判断位置402、403、404の場合を例にして具体的に説明する。まず、それぞれの決定判断位置の直前と直後にある負から正へ変化する零クロス位置411～416を求める。それぞれの決定判断位置について、直前と直後にある零クロス位置411～416の各間隔421、422、423を求める。そ

して、この間隔がピッチ周期とほぼ同じであるものを正弦波に似た波形であると判断する候補とする。つまり、間隔422と間隔423がピッチ周期とほぼ同程度であるので、正弦波に似た波形であると判断する候補とする。

【0027】正弦波に似た波形であると判断する候補としたそれぞれの決定判断位置について、直前と直後の零クロス位置との間にある極大値と極小値の数を求める。決定判断位置403の場合は極大値と極小値はそれぞれ2つつあり、決定判断位置404の場合は極大値と極小値がそれぞれ1つつある。ここで、極大値と極小値が1つつあるもののみを、正弦波に似た波形であると判断し、線形予測残差信号から正確なピッチマークを決定できない位置と判断する。つまり、決定判断位置404がその位置になる。

【0028】このように、音声波形の形状から、線形予測残差信号から正確なピッチマークを決定できるのかどうかの判断を行えば、線形予測残差信号からでは判断しにくい場合においても適切に区間を判断、分類することができるので、それぞれ適切な方法でより正確なピッチマークを決定することができる。

【0029】図5(a)、(b)は本発明の実施の形態4における線形予測残差信号のピッチマークを決定する区間を分類する方法の概念を示し、図5(a)は補間処理によりピッチマークを決定する区間が極端に短い場合のピッチマークの決定、図5(b)は補間処理によりピッチマークを決定する区間に最小値を設けた場合のピッチマークの決定の様子を示した図である。図5(a)、(b)において、501、503は線形予測残差信号、502、504はピッチ基本波である。また、図5(a)、(b)の線形予測残差信号501、502およびピッチ基本波502、504はそれぞれ全く同じ波形である。

【0030】図5(a)に示す場合のように、決定可区間5051の最後のピッチマーク514とその直前の零クロス位置の時間間隔 α と、決定可区間5053の最初のピッチマーク516とその直前の零クロス位置との時間間隔 β との差が、ピッチ基本波502のピッチ周期の2割程度よりも大きくなる場合で、補間処理によりピッチマークを決定する補間決定区間5052の間隔がピッチ周期の2~3倍程度と言うように極端に短いと、ピッチマークの間隔が急激に変化してしまう。したがって、このようにして求めたピッチマークを基準にしてピッチ波形を切り出すと、ピッチ波形の位相が急激に変化してしまうため、再び重畳して音声を作成すると異音が混入してしまう。

【0031】そこで、図5(b)に示す場合のように、補間処理によりピッチマークを決定する補間決定区間5055の区間長に最小値を設けることにより、時間間隔 α と時間間隔 β が大きく異なる場合でもピッチマークの間隔が緩やかに変化するようにピッチマークを決定することができる。したがって、このようにして求めたピッチマークを基準にしてピッチ波形を切り出せば、再び重畳して

音声を作成しても、ピッチ波形の位相が緩やかに変化するので、異音の混入を抑えることができる。

【0032】図6は本発明の実施の形態5におけるピッチ波形切り出しを行うピッチマーク決定方法の概念図を示したものである。図6において、601は音声波形、602は線形予測残差信号、603はピッチ基本波である。この実施の形態5に示す例は、有声区間の終了する部分が線形予測残差信号からは正確なピッチマークを決定することができないと判断された場合を示している。

【0033】まず、線形予測残差信号から正確なピッチマークを決定できる決定可区間6051のピッチマークを前記実施の形態1で説明した従来の方法により決定する。そして、区間6051の最後のピッチマーク6042がその直前にある零クロス位置6061からどれだけ遅れているかを求め、この遅れを時間間隔 α とする。次に、補間処理によりピッチマークを決定する区間6052の全ての零クロス位置6062~6067を求め、この零クロス位置より時間間隔 α だけ直後の位置をそれぞれ補間処理により決定したピッチマークにする。図6におけるピッチマーク6043~6048がこの補間処理により決定したピッチマークである。

【0034】このようにして、ピッチマークを決定すれば、補間処理する区間の両端に線形予測残差信号から決定したピッチマークがない場合でも補間処理により正確なピッチマークを決定することができる。

【0035】なお、前記実施の形態5の説明では、有声区間の最後の部分で補間処理を行う場合を説明したが、有声区間の先頭で補間処理を行う場合にも同様の方法により正確なピッチマークを決定することができる。

【0036】次に、図7は本発明の実施の形態6におけるピッチ波形切り出しを行うピッチマーク決定装置の概略構成を示すブロック図である。図7において、701は有声/無声判定部(U/V判定部)、702は線形予測逆フィルタ(LPC(linear predictive coding)逆フィルタ)、703はピッチフィルタ、704は線形予測残差信号からピッチマークを決定するピッチマーク決定部、705は線形予測残差信号からピッチマークを決定する区間であるのか補間処理によりピッチマークを決定する区間であるのか分類する区間分類部、706は補間処理を行う補間部、707はピッチマークを出力するピッチマークの出力部である。

【0037】まず、音声波形を有声/無声判定部701に投入し、有声区間と判断した区間のみをピッチマークと決定する対象とする。次に、有声区間の音声波形を線形予測逆フィルタ702に投入し、線形予測残差信号を作成する。この線形予測残差信号を区間分類部705に投入する。区間分類部705では、線形予測残差信号から正確なピッチマークを決定できる区間とできない区間の判別を行う。例えば、区間分類部705は判別する位置付近における線形予測残差信号の短区間平均パワーが、ある閾値以上の場合には線形予測残差信号より正確なピッチマーク

を決定できる区間であると判断し、そうでない部分は線形予測残差信号からは正確なピッチマークを決定できない区間と判断する。そして、線形予測残差信号から正確なピッチマークを決定できると判断した区間が有声区間の中でどの部分であるかという区間位置情報をピッチマーク決定部704に出力し、できないと判断した区間が有声区間の中でどの部分であるかという区間位置情報を補間部706に出力する。

【0038】ピッチマーク決定部704では、区間分類部705から出力された区間位置情報をもとに、その区間のピッチマークを従来の方法により決定する。一方、補間部706では、区間分類部705から出力された区間位置情報をもとに、その区間のピッチマークをピッチ基本波とピッチマーク決定部704で決定したピッチマークを利用して線形補間する。ピッチマーク決定部704と補間部706で決定したピッチマークを出力部707に入力し、それぞれのピッチマークを統合して全体のピッチマークを出力する。

【0039】このように、区間分類部705と補間部706を設けることにより、従来では正確なピッチマークを決定することができなかった部分についても正確なピッチマークを決定することができる。

【0040】また、図8は本発明の実施の形態7におけるピッチ波形切り出しを行うピッチマーク決定装置の補間部の概略構成を示すブロック図である。図8において、801はピッチ波数抽出部、802は零クロス位置抽出部、803はピッチマーク線形補間部である。

【0041】ピッチ波数抽出部801は、区間位置情報で与えられる区間内のピッチの波数をピッチ基本波の負から正へ変化する零クロス位置の数を数えることにより抽出する。零クロス位置抽出部802は区間位置情報で与えられる区間内のピッチ基本波が負から正へ変化する零クロス位置を抽出する。ピッチマーク線形補間部803は線形予測残差信号から求めたピッチマークと、ピッチ波数抽出部801で求めたピッチの波数と、零クロス位置抽出部802で求めた零クロス位置を用いてピッチマークを線形補間することにより決定する。

【0042】このようにピッチ波数抽出部801と零クロス位置抽出部802とピッチマーク線形補間部803で補間部を構成すれば、非常に簡単な構成で正確、かつ安定したピッチマークを補間する装置を作ることができる。

【0043】

【発明の効果】以上説明したように、本発明によれば、音声波形から線形予測残差信号を抽出し、これにより正確なピッチマークを決定できる区間とできない区間の判別を行い、正確なピッチマークを決定できると判断された区間のみで短区間変形自己相関係数の最大位置によりピッチマークを決定し、それ以外の区間は音声波形をピッチフィルタにより求めたピッチ基本波を元にその零クロス位置と波数からピッチマークを決定するため、非常

に少ない演算量、かつ装置構成で正確なピッチマークを決定することができ、決定したピッチマークを用いて、ピッチ波形を抽出し、再び重畳して音声合成すれば、極めて高品質の合成音声を得ることができるという効果を奏する。

【図面の簡単な説明】

【図1】本発明の実施の形態1におけるピッチ波形切り出しを行うピッチマーク決定方法の概念図を示したものである。

【図2】本実施の形態1におけるピッチ基本波を利用したピッチマークの補間方法の概念図を示したものである。

【図3】本発明の実施の形態2における線形予測残差信号のピッチマークを決定する区間を分類する方法の概念図を示したものである。

【図4】本発明の実施の形態3における線形予測残差信号のピッチマークを決定する区間を分類する方法の概念図を示したものである。

【図5】本発明の実施の形態4における線形予測残差信号のピッチマークを決定する区間を分類する方法の概念図を示し、(a)は補間処理によりピッチマークを決定する区間が極端に短い場合のピッチマークの決定、(b)は補間処理によりピッチマークを決定する区間に最小値を設けた場合のピッチマークの決定の様子を示した図である。

【図6】本発明の実施の形態5におけるピッチ波形切り出しを行うピッチマーク決定方法の概念図を示したものである。

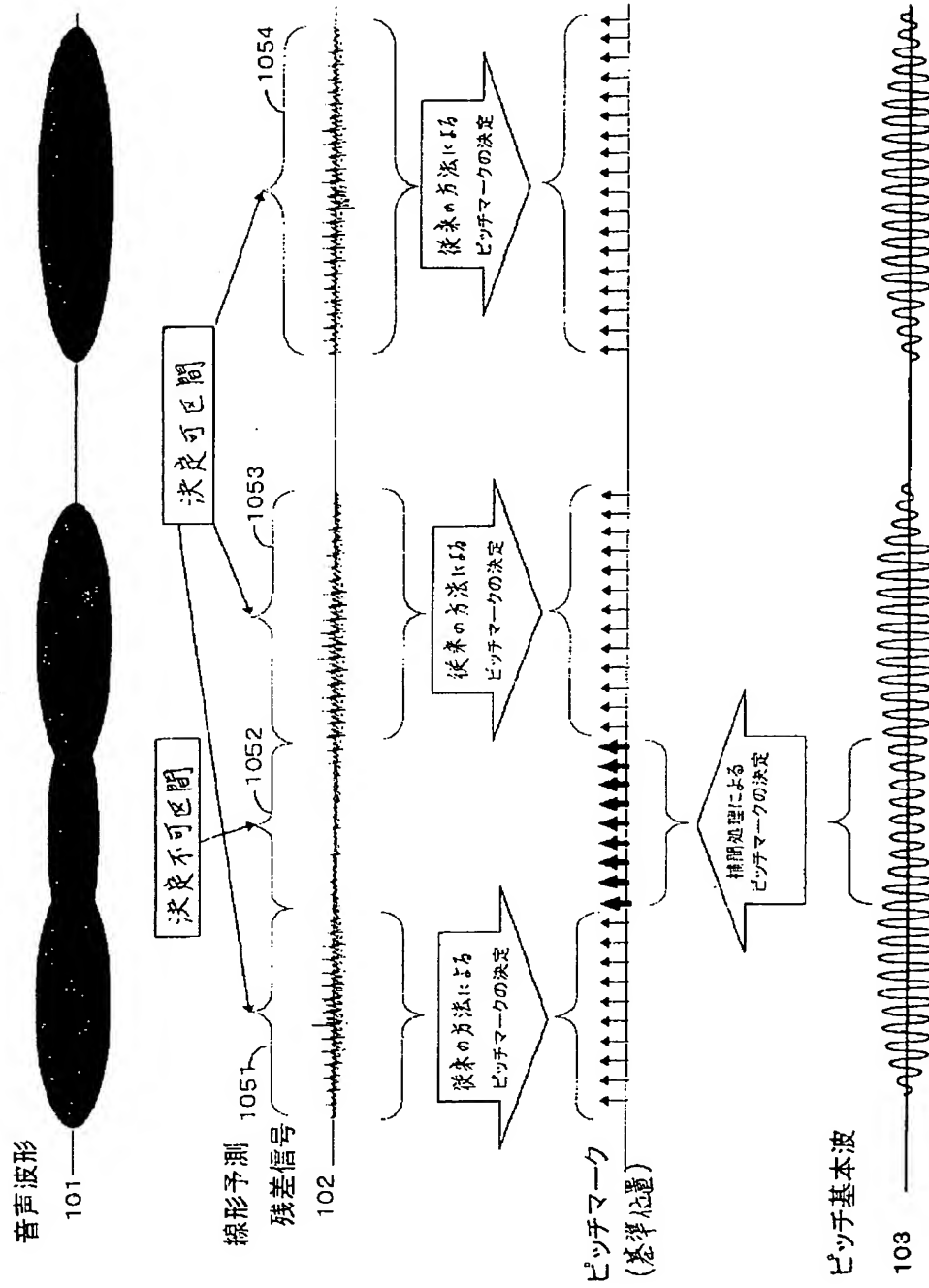
【図7】本発明の実施の形態6におけるピッチ波形切り出しを行うピッチマーク決定装置の概略構成を示すブロック図である。

【図8】本発明の実施の形態7におけるピッチ波形切り出しを行うピッチマーク決定装置の補間部の概略構成を示すブロック図である。

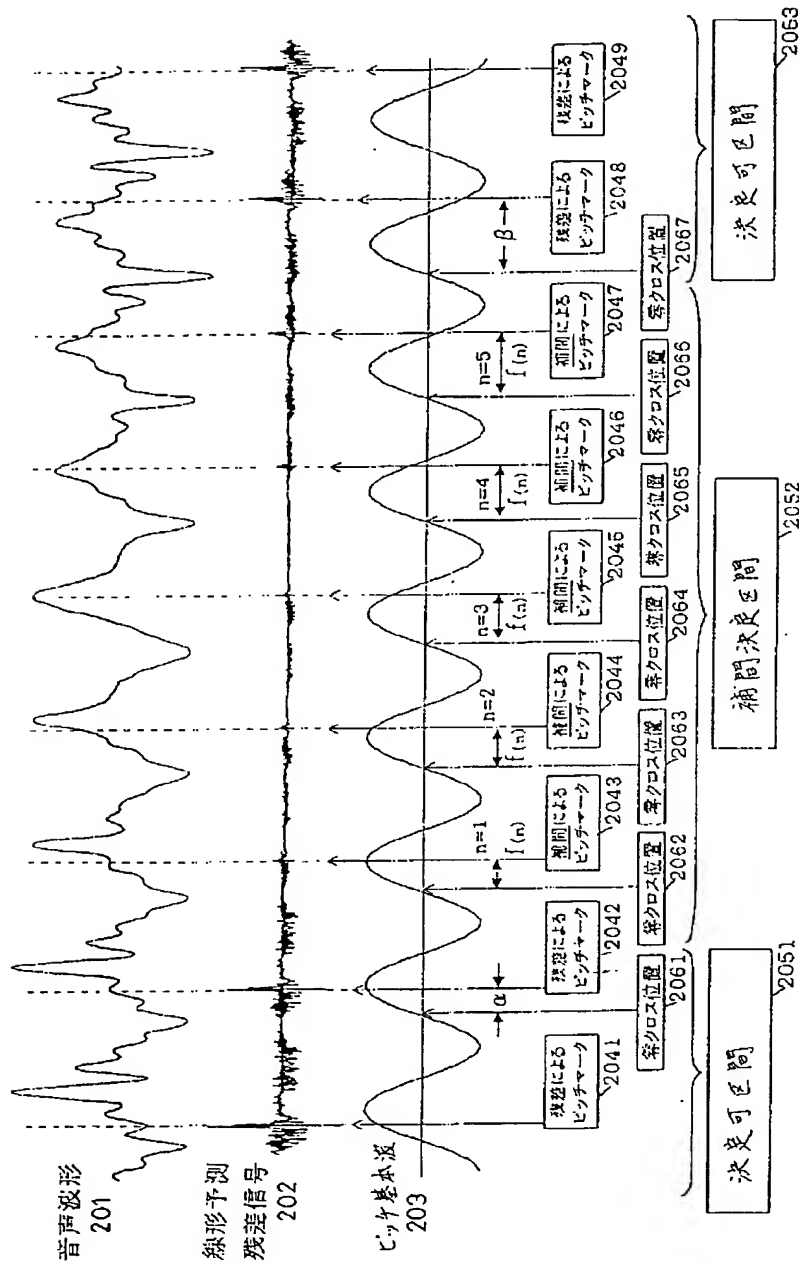
【符号の説明】

101, 201, 401, 601…音声波形、 102, 202, 301, 501, 503, 602…線形予測残差信号、 103, 203, 502, 504, 603…ピッチ基本波、 302, 303, 402, 403, 404…決定判断位置、 304…広範囲の最大値、 305…閾値、 411, 413, 415…直前の零クロス位置、 412, 414, 416…直後の零クロス位置、 421, 422, 423…間隔、 510～529, 2041～2049, 6041～6048…ピッチマーク、 701…有声/無声判定部(U/V判定部)、 702…線形予測逆フィルタ、 703…ピッチフィルタ、 704…ピッチマーク決定部、 705…区間分類部、 706…補間部、 707…出力部、 801…ピッチ波数抽出部、 802…零クロス位置抽出部、 803…ピッチマーク線形補間部、 1051, 1053, 1054, 2051, 2053, 5051, 5053, 5054, 5056, 6051…決定可区間、 1052, 2052, …決定不可区間、 2061～2067, 6061～6067…零クロス位置。

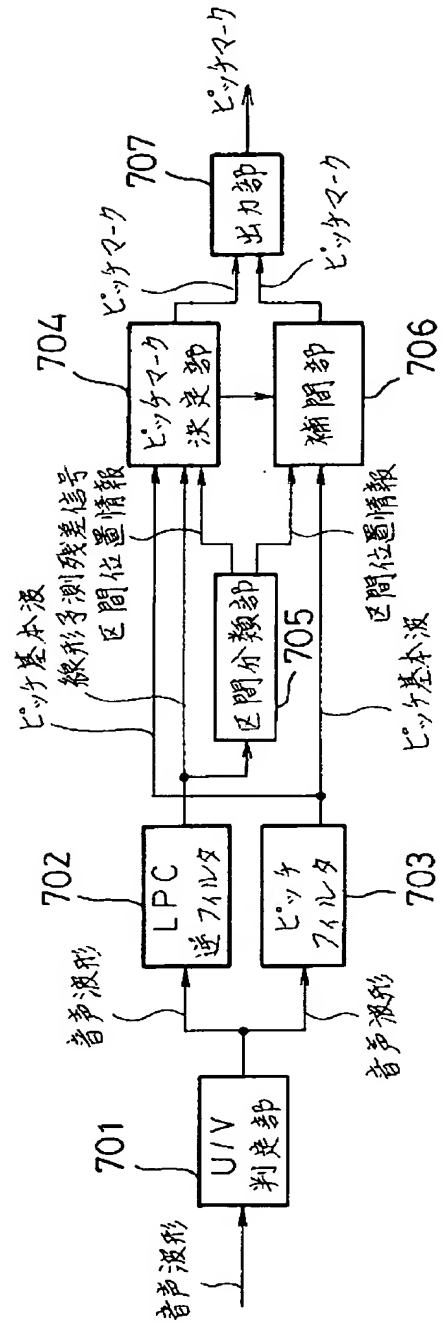
【図1】



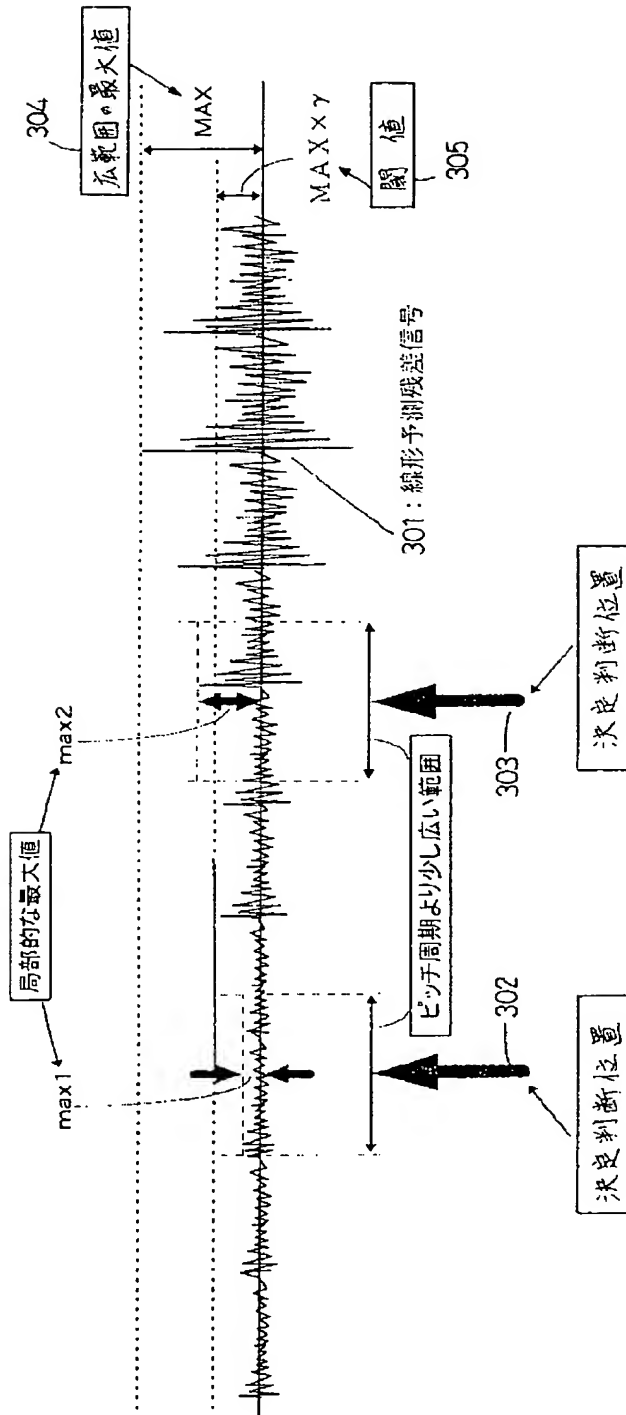
【図2】



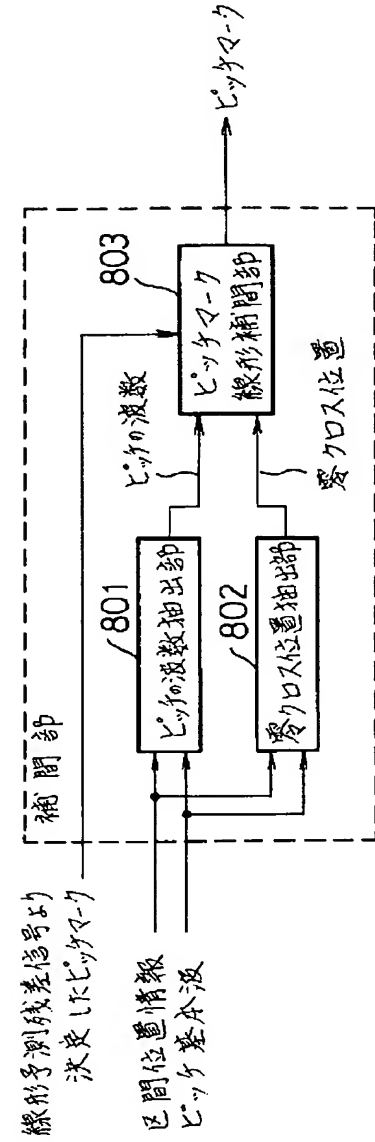
【図7】



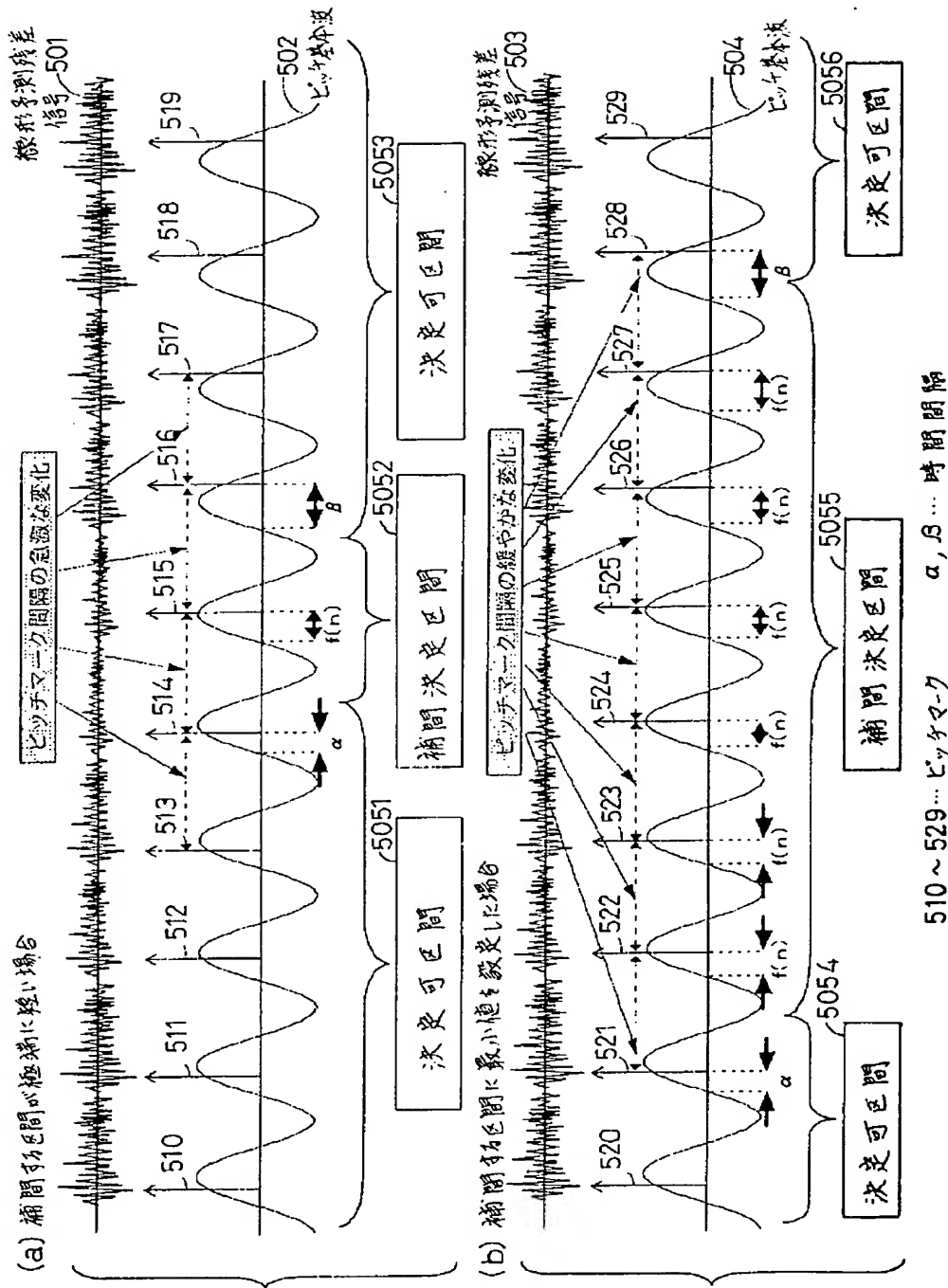
【図3】



【図8】



【図5】



【図6】

